JOURNAL OF BACTERIOLOGY, Oct. 1986, p. 40–48
0021-9193/86/100040-09$02.00/0

Vol. 168, No. 1

# Organization and Expression of Genetic Determinants for Synthesis and Assembly of Type 51 R Bodies

JOSEPH A. KANABROCKI,† ROBERT L. QUACKENBUSH,* AND FINN R. POND

*Department of Microbiology, School of Medicine, University of South Dakota, Vermillion, South Dakota 57069*

Received 19 February 1986/Accepted 11 June 1986

**Type 51 R bodies are produced by all bacterial endosymbionts (*Caedibacter taeniospiralis*) of *Paramecium tetraurelia* that confer the hump-killer trait upon their hosts. Type 51 R-body synthesis by *C. taeniospiralis* is required for expression of the hump-killer trait. The genetic determinants for type 51 R-body synthesis by *C. taeniospiralis* 47 have been cloned and expressed in *Escherichia coli*. In this communication we describe three species of polypeptides required for R-body synthesis and the organization of their genetic determinants. Each polypeptide species is controlled by a separate gene that is expressed as an independent transcriptional unit possessing regulatory signals that are recognized by *E. coli*. Two polypeptide species of 10 and 18 kilodaltons are required for R-body synthesis but apparently are not structural subunits. The third polypeptide species (13 kilodaltons) is the major structural subunit. R-body assembly involves polymerization reactions that result in high-molecular-mass polypeptide complexes, primarily composed of the 13-kilodalton polypeptide species, that appear to be the result of covalent cross-linking between structural subunits. The results presented here have been suggested to apply to the assembly and structure of all type 51 R bodies, but not necessarily to other R-body types.**

Bacterial endosymbionts commonly occur in various species of *Paramecium*. The most notable group of these bacteria comprise the genus *Caedibacter*, more commonly known as kappa or kappa particles. These bacteria are easily distinguished from other symbionts by their ability to produce refractile inclusion bodies known as R bodies. An R body is a proteinaceous ribbon (approximately 10 to 20 μm long, 0.5 μm wide, and 13 nm thick) that is rolled up inside the bacterial cell, appearing as a hollow cylindical structure (1) (Fig. 1). Within any given population of *Caedibacter* species only a fraction of the individuals possess R bodies. R body-containing forms of *Caedibacter* species are toxic to sensitive strains of paramecia. Thus, paramecia that carry any *Caedibacter* strain are referred to as killers. Killer paramecia apparently release some of their endosymbionts into the environment when food vacuoles are emptied via the cytopyge (15). Killing occurs after an R body-containing symbiont is ingested by a sensitive paramecium.

In three of the four species of *Caedibacter*, R-body synthesis appears to be determined by extrachromosomal elements that are thought to be defective bacteriophages (15). In the fourth species, *Caedibacter taeniospiralis*, R-body synthesis is determined by plasmids (17). R bodies produced by *C. taeniospiralis*, type 51 R bodies, are distinguished from those produced by the other species of *Caedibacter* on the basis of morphology and their response to changes in pH. Type 51 R-body ribbons typically end in acute angles and unroll from the inside, in a telescoping fashion, when the pH is lowered to 6.5 or less. Loss of the ability to assemble type 51 R bodies by *C. taeniospiralis* is accompanied by loss of toxic activity against sensitive strains of paramecia (4, 23). The genetic determinants for type 51 R-body synthesis and assembly from the R body-coding plasmid (pKAP47) of *C. taeniospiralis* 47 have been cloned and expressed in *Escherichia coli* (17) (Fig. 1).

Past attempts at studying the structural components of R bodies have been unsuccessful because of two obstacles. First, it was diffficult to obtain sufficient amounts of purified R bodies, since *Caedibacter* species cannot be grown outside of their hosts. Second, earlier attempts to characterize the component parts of R bodies with polyacrylamide gels failed because R bodies are resistant to standard methods for dissociating protein assemblies (14). Cloning and expression of the genetic determinants of type 51 R bodies in *E. coli* has removed the first obstacle, thus lessening the problems presented by the second. We report here that at least three proteins are involved in the structure and assembly of type 51 R bodies. In addition, the approximate locations of the genetic determinants for each of these proteins are presented.

## MATERIALS AND METHODS

**Bacterial strains and plasmids.** The strains of *E. coli* used in this study are 294 (22; *endA hsdR thi pro*) and the minicell-producing strain, P678-54 (9; *thr-1 leuB6 minA1 minB2 thi-1 ara-13 lacY1 gal-6 malA1 xyl-7 mtl-2 azi-8 rpsL135 fhuA2*), and plasmid-containing derivatives thereof. The plasmids used in this study are as follows: the R body-coding plasmid pBQ51 (17) and its derivative, pBQ65, which is described in Results; the cloning vehicle pBR327 (21), which was kindly supplied by Francisco Bolivar (Universidad Nacional Autonoma de Mexico); and deletion derivatives of pBQ65, also described in Results.

**Bacterial growth.** Minicell-producing strains were grown, for the purpose of minicell purification, in M9 minimal medium (0.1 mM $CaCl_2$, 1 mM $MgSO_4$, 0.6% $Na_2HPO_4$, 0.3% $KH_2PO_4$, 0.05% NaCl, 0.1% $NH_4Cl$) supplemented with 0.4% glucose, 1% Casamino Acids (Difco Laboratories, Detroit, Mich.), 4 mg of thiamine hydrochloride per liter, 80 mg of threonine per liter, 40 mg of leucine per liter, and 200

* Corresponding author.
† Present address: Department of Genetics and Development, Cornell University, Ithaca, NY 14850.

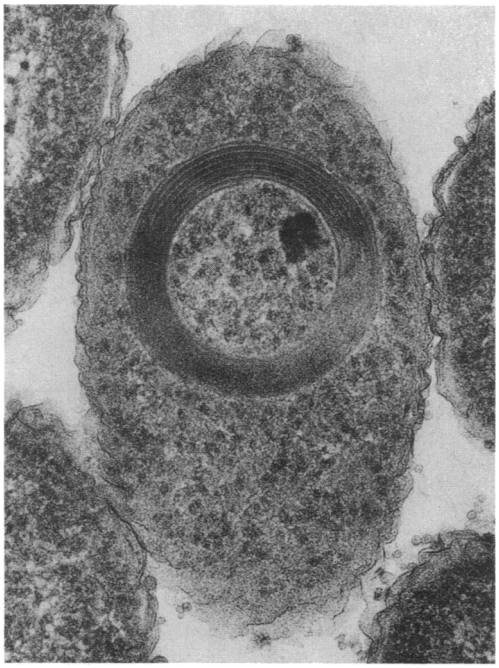

FIG. 1. Electron micrograph of a grazing section of *E. coli* 294(pBQ51) containing an R body (cross section) ×68,400.

mg of streptomycin per liter. Otherwise, all strains were grown in L broth (0.5% NaCl, 1% tryptone, and 0.5% yeast extract). All cultures were grown with shaking at 37°C.

**Purification of R bodies.** R bodies were purified from bacterial lysates by differential centrifugation or velocity centrifugation through sucrose gradients. Lysates were prepared as follows. Cell pellets, obtained by centrifugation (4,420 × *g* for 15 min) of 250 ml of late-stationary-phase cultures, were suspended in 5 ml of 25% sucrose mix (in 50 mM Tris hydrochloride, pH 8.0). Lysozyme was added to a final concentration of 2.5 mg/ml and incubated on ice for 15 min. DNaseI and RNase were then added to a final concentration of 2 mg/ml, and the mixture was held on ice for an additional 30 min. A 15-ml sample of 2% sodium lauryl sarcosinate in 50 mM Tris (pH 8.0), was then added to the cell suspension with mixing, and the cells were held on ice until significant lysis was achieved.

R bodies were purified from lysates of *E. coli* 294(pBQ51) by differential centrifugation with the following schedule of centrifugation steps: 7,750 × *g* for 5 s, 7,750 × *g* for 1 min, 7,750 × *g* for 2 min, 7,750 × *g* for 4 min, 12,100 × *g* for 2.5 min, 12,100 × *g* for 5 min, 12,100 × *g* for 10 min, and 31,000 × *g* for 10 min. Pellets formed during each centrifugation step were collected, suspended in deionized $H_2O$, and stored at 4°C. The supernatant fluid was transferred to a fresh centrifuge tube and subjected to the next step on the centrifugation schedule.

R bodies were purified from lysates of *E. coli* P678-54(pBQ65) by using sucrose gradients. Pellets obtained by centrifugation of lysates at 31,000 × *g* for 20 min were suspended in 50 mM Tris hydrochloride (pH 8.0) and layered on 25-ml linear sucrose gradients (10 to 55% [wt/vol] in 50 mM Tris hydrochloride, pH 8.0). The gradients were then subjected to centrifugation at 1,500 × *g* for 45 min. The resulting R body-containing bands were removed from the middle of the gradients with a Pasteur pipette and pelleted by centrifugation at 31,000 × *g* for 15 min. The R-body pellets were then suspended in 1 ml of 50 mM Tris hydrochloride (pH 8.0) and passed through two more sucrose gradients, repeating the steps described above. Purified R-body pellets were stored at −20°C until needed.

**Plasmid purification and manipulations.** Plasmid DNA was prepared by the cleared lysate method of Clewell and Helinski (3). The procedures used for restriction endonuclease analysis, cloning, and transformation have been described previously (17). Recombinant plasmids were detected with the miniscreening protocol of Birnboim and Doly (2). Restriction endonucleases and T4 DNA ligase were purchased from New England Biolabs (Beverly, Mass.) and Bethesda Research Laboratories (Gaithersburg, Md.), respectively.

**Polyacrylamide gel electrophoresis of proteins.** Sodium dodecyl sulfate (SDS)-polyacrylamide gels were prepared as described by Laemmli (6). Unless otherwise specified, samples were suspended in final sample buffer (62.5 mM Tris hydrochloride [pH 6.8], 2% SDS, 10% glycerol, 5% 2-mercaptoethanol, and 0.001% bromphenol blue), treated as described by Laemmli (6), and applied to SDS-polyacrylamide gels. Two-dimensional gel electrophoresis was performed as described by O'Farrell (11). Gels to be stained were placed in staining solution (10% glacial acetic acid, 45% methanol, 0.025% Coomassie blue R-250) overnight with gentle shaking and destained with multiple changes of destaining solution (10% glacial acetic acid, 45% methanol) with gentle shaking. Gels containing radiolabeled samples were treated for fluorography with Enlightning (New England Nuclear Corp., Boston, Mass.) according to the manufacturer's instructions and then autoradiographed at −70°C with Kodak SB5 film.

**In vitro transcription-translation system.** In vitro protein synthesis was performed with an S30 system prepared as described by Pratt (13). Reactions were carried out in 150-μl volumes prepared by mixing 25 μl of plasmid (5 μg in deionized water) with 75 μl of reaction mix (4 mM Tris acetate [pH 8.2], 55 mM potassium acetate, 27 mM ammonium acetate, 1.4 mM UTP, 0.5 mM CTP, 0.2 mM each amino acid, 21 mM phosphoenolpyruvate, 2.7 μg of *p*-amino benzoic acid per ml, 2.7 μg of NADP per ml, 2.7 μg of pyridoxine per ml, 2.7 μg of FAD per ml, 7.4 mM calcium chloride) and 50 μl of S30 to which 30 μCi of [$^{35}$S]methionine (New England Nuclear) had been added. Cell-free synthesis was allowed to proceed for 60 min at 37°C and then terminated by adding 100 μl of 5% trichloroacetic acid. Acid-insoluble molecules were pelleted by centrifugation at 13,000 × *g* for 15 min.

**Radiolabeling of plasmid-encoded polypeptides in minicells.** Minicells produced by *E. coli* P678-54 and its derivatives were purified on linear sucrose gradients as described by Roozen et al. (19). Purified minicells were suspended in 2 ml of methionine assay medium (Difco) and dispensed into 0.5-ml working samples. After incubation at 37°C for 15 min, 20 μCi of [$^{35}$S]methionine was added to each suspension and then incubated for the desired time of labeling. Minicells were removed from unabsorbed label by centrifugation (13,000 × *g* for 5 min) followed by three washes in 67 mM phosphate buffer (pH 6.8). In pulse-chase experiments, minicells were suspended in supplemented M9 medium and incubated at 37°C for the desired chase period. Protein synthesis was terminated at designated times by the addition of 100 μl of 5% trichloroacetic acid. The samples were then pelleted (13,000 × *g* for 5 min), washed three times in 67 mM phosphate buffer (pH 6.8), and stored at −70°C.

## RESULTS

**Construction of pBQ65.** In a previous report (17) we demonstrated, by analysis of the expression of deletion derivatives of a recombinant plasmid (pBQ51) containing the R body-coding sequence from pKAP47 (the R body-coding plasmid of *C. taeniospiralis* 47) in *E. coli*, that the total DNA sequence coding for R body synthesis contains less than 2,800 base pairs (17). To reduce the amount of sequence not essential for R-body synthesis, a 2,700-base-pair *Sph*I-*Cla*I fragment from that region of pBQ51 was inserted in pBR327. The resulting plasmid, pBQ65 (Fig. 2), retains the genetic determinants for R-body synthesis. However, the *Sph*I recognition sequence apparently was damaged during the cloning process and is not present in pBQ65.

**Polypeptides associated with purified R bodies.** R bodies obtained from *E. coli* 294(pBQ51), *E. coli* 294(pBQ65), or *E. coli* P678-54(pBQ65) were not completely dissociated when treated with SDS (up to 10%) and 5% 2-mercaptoethanol with boiling (up to 1 h). However, boiling in final sample buffer for 5 min caused at least some of the R body-associated polypeptides to dissociate from the R-body structures. More than 30 polypeptide bands, ranging from less than 14 to over 1,000 kilodaltons (kDa), dissociated from purified R bodies could be resolved on SDS-polyacrylamide gels (Fig. 3). The pattern of polypeptide bands generated was the same regardless of the source of the R bodies. The most notable feature of the pattern generated by SDS-polyacrylamide electrophoresis is a ladderlike profile of polypeptides starting at about 40 kDa and extending to the top of each gel. In fact, these samples contain material that does not even enter the running gel from the stacking gel. Assuming that these polypeptides are coded for by the insert contained in pBQ65, it is not possible that each of them represents a separate gene product because of the lack of sufficient coding capacity. The ladderlike profile strongly suggests that the higher-molecular-mass (>40-kDa) polypeptide species are polymerization products of low-molecular-mass polypeptides and thus may represent intermediates in R-body assembly. Although the intervals between polypeptides in the ladder appear to be regular, no constant interval or predictable pattern could be detected from the polypeptide molecular weights (as estimated from standard curves derived from the migration of molecular weight standards on the gels). In addition, since the percentage of acrylamide in the gels was varied, the migration of ladder polypeptides was inconsistent relative to the migration of molecular weight standards (Fig. 3). With an increased percentage of acrylamide, the migration of ladder polypeptides decreased relative to that of the molecular weight standards. This indicates that the use of a standard curve determined by the migration of molecular weight standards may not be valid for determining the sizes of these polypeptides. Thus when we refer to polypeptides in terms of kilodaltons we are not making any claims with respect to actual molecular mass, only to the relative migration of the polypeptide through an SDS-polyacrylamide gel with respect to molecular weight standards.

If the ladder polypeptides are oligomers of low-molecular-weight subunits, the question of the nature of the intermolecular interactions involved in the assembly of the ladder polypeptides arises. The following denaturing treatments had no effect on the pattern of ladder polypeptides generated by polyacrylamide gel electrophoresis: 8 M urea; any combination of SDS (up to 10%), 2-mercaptoethanol, dithiothreitol, and boiling (up to 1 h); boiling in dilute HCl (pH 1.8) for 5 min followed by standard treatment in final sample buffer; incubation for 2 h in 8.6 M guanidine hydrochloride at 37°C followed by dialysis and standard treatment

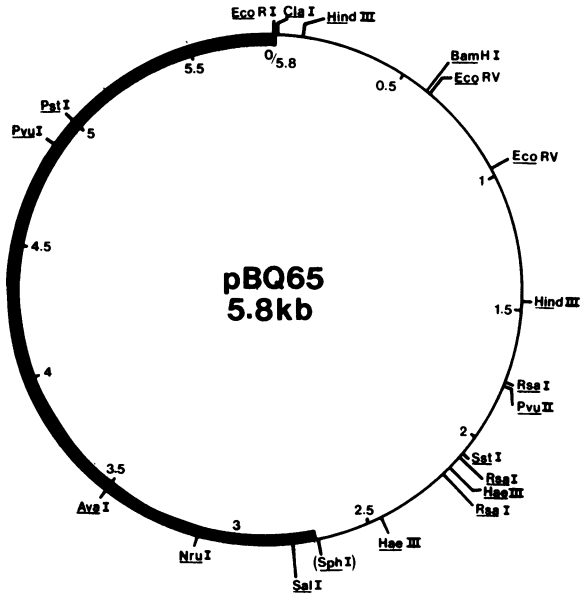


FIG. 2. Restriction map of pBQ65. Sequences derived from cloning vehicle pBR327 are designated by the thickened portion of the circle. The *Rsa*I and *Hae*III sites present in the cloning vehicle-derived portion of pBQ65 are not shown. kb, kilobases.

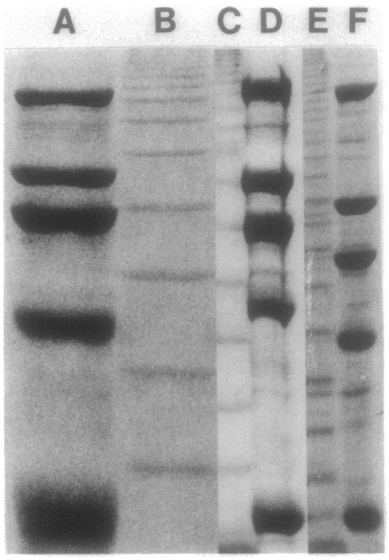


FIG. 3. Proteins associated with purified R bodies. The photograph is a composite of three separate polyacrylamide gels differing in percentage of acrylamide (lanes A and B, 7 to 20% gradient; lanes C and D, 7%; and lanes E and F, 5%). The magnification of each gel photograph was adjusted so that the distances between the largest and smallest molecular mass standards were the same for each. Lanes: A, D, and F, molecular weight standards (myosin, 200,000 Da; β-galactosidase, 116,250 Da; phosphorylase B, 66,200 Da; bovine serum albumin, 66,200 Da; and ovalbumin, 45,000 Da); B and C, purified R bodies from *E. coli* P678-54(pBQ65); E, purified R bodies from *E. coli* 294(pBQ51). All samples were boiled for 5 min in final sample buffer before application to gels. All gels were stained with Coomassie blue R-250.

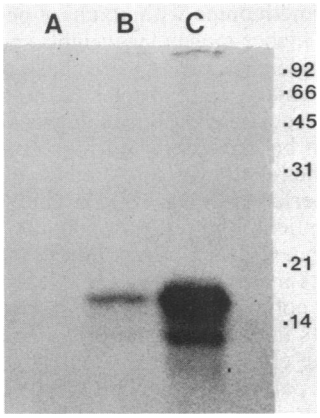


FIG. 4. pBQ65 in vitro translation products resolved by electrophoresis on a 17% SDS–polyacrylamide gel. [$^{35}$S]methionine-labeled polypeptides were synthesized in the coupled transcription-translation system as described in the text. Shown is the fluorograph of the gel. Lanes: A, pBQ59, a deletion derivative of pBQ51 unable to direct R body synthesis; B, pBQ65; C, control. Molecular weight standards: phosphorylase B, 92,500 Da; bovine serum albumin, 66,200 Da; ovalbumin, 45,000 Da; carbonic anhydrase, 31,000 Da; soybean trypsin inhibitor, 21,500 Da; and lysozyme, 14,400 Da.

with final sample buffer; and incubation in 6 M guanidine thiocyanate at 37°C followed by dialysis and standard treatment with final sample buffer. These results demonstrate that it is highly unlikely that the subunits of the ladder polypeptides are solely held together by disulfide bridges, ionic or hydrogen bonds, or hydrophobic interactions. If the ladder polypeptides are polymers of low-molecular-weight

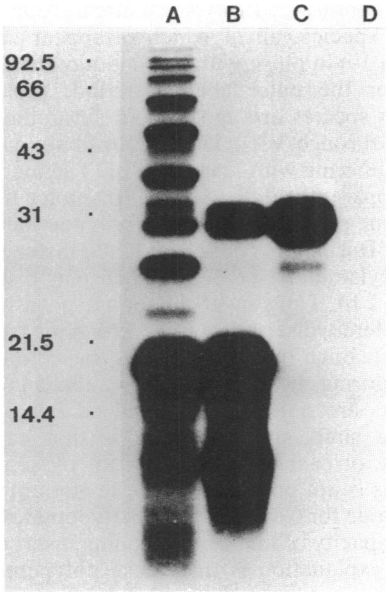


FIG. 5. Selective expression of plasmid-encoded polypeptides in minicells. Purified minicells were incubated in the presence of [$^{35}$S]methionine for 60 min with no chase period. Samples were then boiled in final sample buffer for 5 min. Polypeptides were resolved by electrophoresis on a 17% SDS–polyacrylamide gel. Shown is the fluorograph of the gel. Lanes: A, *E. coli* P678-54(pBQ65); B, *E. coli* P678-54(pBQ59); C, *E. coli* P678-54(pBR327); D, *E. coli* P678-54. Molecular weight standards are the same as in Fig. 4.

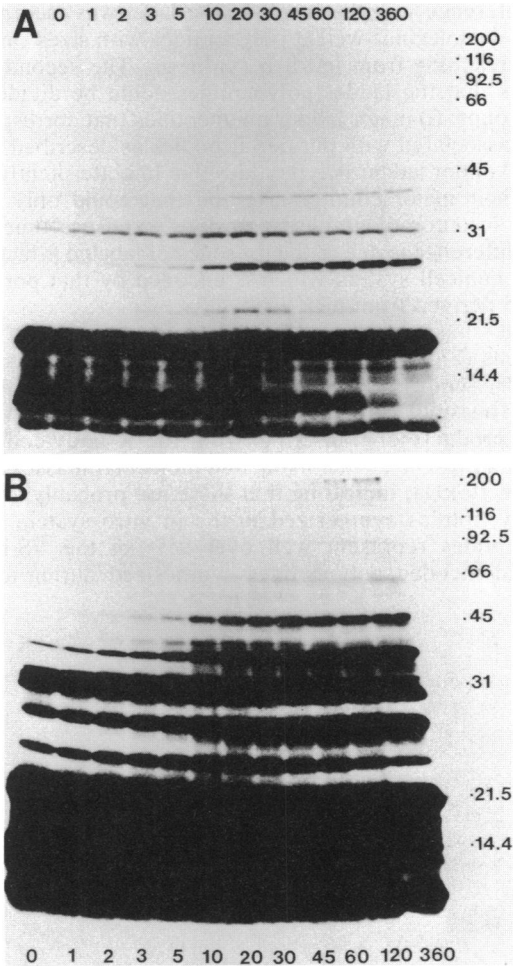


FIG. 6. Pulse-chase analysis of polypeptide products resulting from expression of the R body-coding sequence of pBQ65. Minicells purified from *E. coli* P678-54(pBQ65) were labeled in the presence of [$^{35}$S]methionine for 3 min and chased with unlabeled methionine for the times specified (in minutes). The samples were boiled in final sample buffer for 5 min and were resolved by electrophoresis on a 7 to 20% SDS–polyacrylamide gradient gel. Shown are fluorographs of the gel; A, 12-h exposure; B, 144-h exposure.

subunits, these results strongly indicate that the polymerization process involves covalent cross-linking of subunits.

**Characterization of polypeptides encoded by the genetic determinants of type 51 R bodies.** To determine what polypeptides are actually coded for by the cloned R body-coding DNA sequence, expression of pBQ65 in vitro and in minicells was studied. The addition of pBQ65 to the coupled transcription-translation system as described above resulted in the synthesis of three polypeptides that migrated at positions in SDS-polyacrylamide gels corresponding to 18, 13, and 10 kDa (Fig. 4). No other polypeptide products were detected. Assuming that the in vitro system correctly responds to regulatory signals in the R body-coding sequence, these results indicate that at least three low-molecular-weight polypeptides are coded for by this DNA sequence.

In contrast to results obtained with the in vitro system, the pattern of pBQ65-encoded polypeptides from minicells generated by electrophoresis on SDS-polyacrylamide gels was similar to that obtained from purified R bodies (Fig. 5). However, several differences were noted. The most appar-

ent difference was that most of the label was incorporated into low-molecular-weight polypeptides with sizes similar to those resulting from in vitro synthesis. The second difference is that the ladder polypeptides could be divided into two groups: (i) major ladder polypeptides that correspond to those associated with purified R bodies as described above; and (ii) minor ladder polypeptides that migrate slightly faster than their major counterparts and that could only be detected in autoradiographs given long exposure times. The third difference is due to the presence of labeled β-lactamase in the minicell system which is encoded by that portion of pBQ65 derived from pBR327.

Pulse-chase studies of the expression of pBQ65 in minicells (Fig. 6) demonstrate that ladder polypeptides are indeed assembly products of low-molecular-weight subunits. Three (possibly four) polypeptides appear to be the translational products of transcripts from the R body-coding sequence of pBQ65. Their estimated molecular masses are 18, 13, and 10 kDa, indicating that these are probably identical to polypeptides synthesized in the in vitro system. These polypeptides represent well over 95% of the $^{35}S$-labeled plasmid-encoded polypeptides synthesized during a 3-min pulse with [$^{35}S$]methionine with no chase period. During the chase period, higher-molecular-weight species, primarily ladder polypeptides, appear. As the chase period proceeds, the higher-molecular-weight species become more apparent (Fig. 6). The polypeptides produced during a 3-min pulse and a 6-h chase can be separated into four groups. In group 1 polypeptides, the relative amounts remain constant throughout the chase period (10 kDa, 18 kDa, and β-lactamase). In group 2 polypeptides, the relative amounts decrease during the chase period (13 kDa). In group 3 polypeptides, the relative amounts increase during the chase period (27 kDa, 40 kDa, ladder polypeptides). In group 4 polypeptides, the relative amounts initially rise and then diminish about halfway through the chase period (24 and 41 kDa). It is likely that the 13-kDa polypeptide which represents group 2 is the structural subunit of R bodies and that the ladder polypeptides (group 3, starting at 27 kDa) are oligomers of the 13-kDa polypeptide. The autoradiograph shown in Fig. 6 does not show all of the group 3 (ladder) polypeptides; polymerization products that are too large to migrate through the stacking gel could be visualized in autoradiographs of stacking gels starting from 30 min into the chase period. Since the 18- and 10-kDa polypeptide species (group 1) are not used up during the chase period, it is most likely that they are required for the assembly of structural subunits and for subunit incorporation into R-body structure but are not an integral part of the R-body structure. The significance of the group 4 polypeptides is not clear. Perhaps they are intermediates formed between structural subunit(s) and catalytic (group 1) polypeptides. However, if group 2 polypeptides are essentially enzyme-substrate complexes, the question that arises is, why are these complexes stable in denaturing and reducing environments?

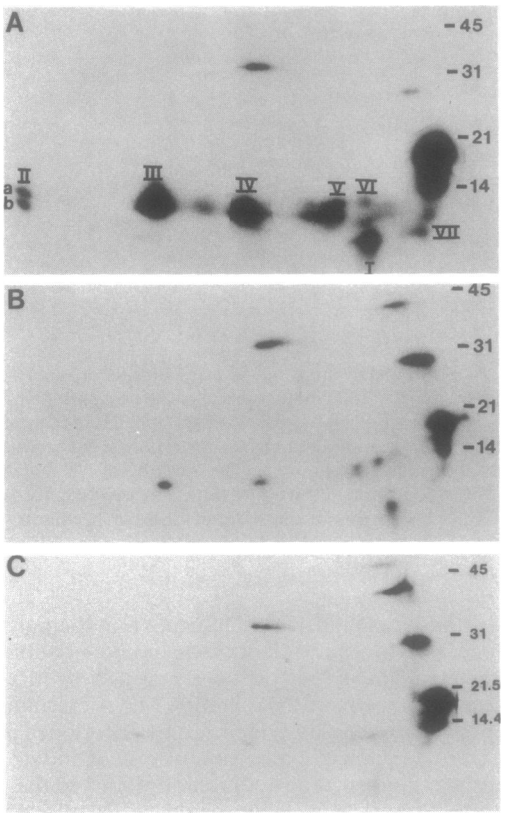


FIG. 7. Two-dimensional gel electrophoresis of polypeptide products resulting from expression of the R body-coding sequences of pBQ65. Purified minicells from *E. coli* P678-54(pBQ65) were pulse-labeled with [$^{35}S$]methionine for 1 min and were chased with unlabeled methionine for 2 min (A), 2 h (B), or 12 h (C). Samples were initially resolved by isoelectric focusing over a pH gradient of 3 to 10 (right to left) and then further separated by electrophoresis on a 7 to 20% SDS–polyacrylamide gradient gel. The fluorographs shown above were obtained from 12-h exposures of the dried gels. The polypeptide designated I in panel A is believed to be the 10-kDa polypeptide, and the polypeptides designated IIa, IIb, and III through VII are believed to be different forms of the 13-kDa polypeptide.

Two-dimensional gels were used in an attempt to determine whether there may be more than one pBQ65-encoded polypeptide species sharing a given molecular mass. The group 3 polypeptide (13 kDa) is actually a group of at least 12 polypeptide species, all of which disappear during a 12-h chase after a 1-min pulse with [$^{35}S$]methionine (Fig. 7). Two minutes after the pulse-labeling period, the 12 group 3 polypeptides species are apparent in autoradiographs (Fig. 7A, spots II through VII). They occur in six paired groups, each group differing with respect to pI. The members within each pair apparently differ slightly in molecular mass but share the same pI. The pIs of these polypeptides range from about 4 to 9. During the chase period the faster-migrating (in SDS-polyacrylamide) member of each pair disappears first (in less than 2 h). The slower-migrating species in each pair almost completely disappear within 12 h after the application of pulse label. Each of the six groups (based on pI) appears to diminish during the chase period at about the same rate. These results are quite unusual and cannot be easily explained. The simplest explanation is that each of these polypeptides, or each pair, is coded for by separate genes. However this is not possible, since the amount of sequence required to code for 6 or 12 of these polypeptides far exceeds the coding capacity of the R body-coding insert of pBQ65. A more likely explanation is that these polypeptides are encoded by one or two genes and exhibit differences in pI, and possibly migration rate, due to posttranslational processing which may be mediated by the 10-kDa or the 18-kDa polypeptides. Furthermore, careful examination of the 18-kDa polypeptide reveals that more than one polypeptide species is present (Fig. 6 and 7). These polypeptides share the same pI (Fig. 7) and are probably the result of posttranslational processing of a common polypeptide precursor.

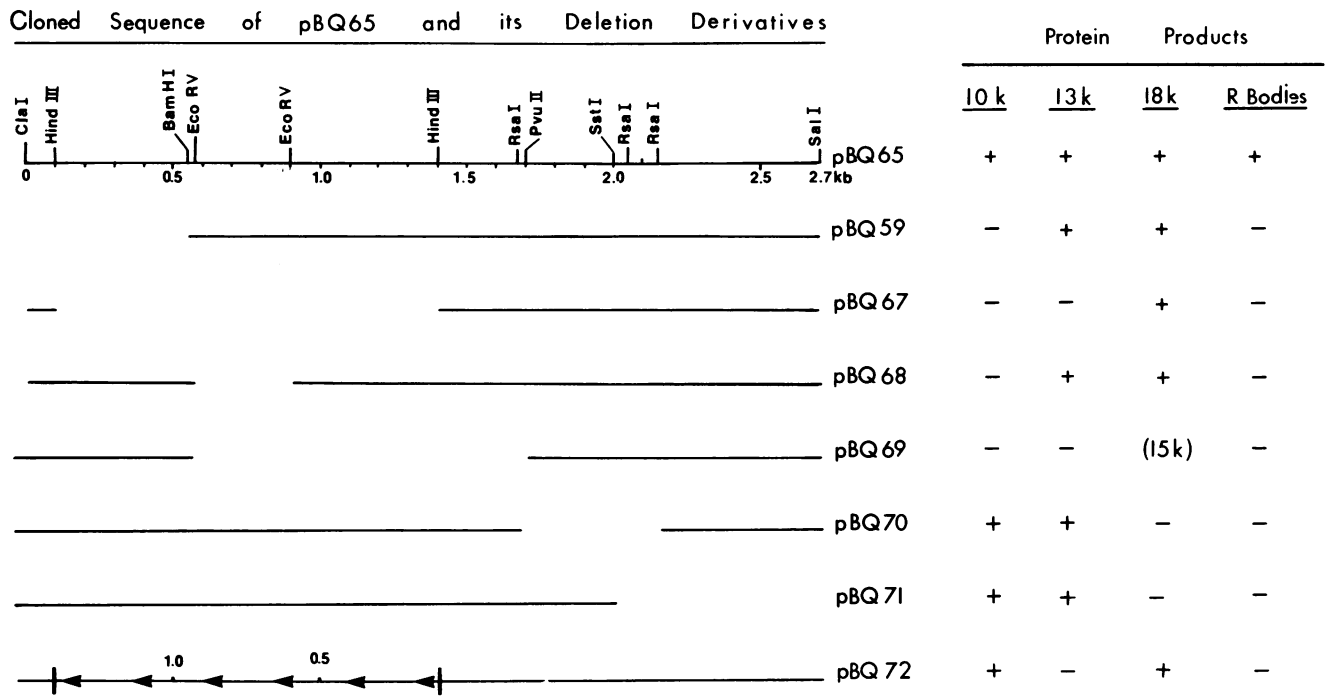


FIG. 8. Physical maps of the cloned sequences of pBQ65 and its deletion derivatives. Maps are shown with deleted regions appearing as gaps. The one exception is pBQ72 which is the result of inverting the 1.3-kilobase-pair *Hin*dIII fragment as indicated by the arrows above. Also included is a summary of the major polypeptide products encoded by each of these sequences as shown in Fig. 9.

**Localization of gene sequences required for synthesis of R bodies.** To determine the approximate location and order of the genes required for expression of R-body synthesis, deletion derivatives of pBQ65 were generated and introduced into *E. coli* P678-54. The resulting plasmids were mapped by restriction endonuclease analysis (Fig. 8), and the polypeptides that they encode were determined by minicell analysis (Fig. 8 and 9). The results demonstrate that R-body synthesis is controlled by three independent transcriptional units occurring in the following order: (i) gene sequence coding for the 18-kDa polypeptide(s), (ii) gene sequence coding for the 13-kDa polypeptides, and (iii) gene sequence coding for the 10-kDa polypeptide. R bodies are not produced when any of these genes are not normally expressed. Thus each of these gene sequences is essential. The locations and maximum limits of the sequences containing these transcriptional units are shown in Fig. 10. These results have been confirmed by DNA sequence data (D. Heruth and R. L. Quackenbush, unpublished observations).

Analysis of the deletion derivatives of pBQ65 and their expression in minicells has also allowed us to determine the approximate location of the regulatory sequence of the gene that encodes the 18-kDa polypeptide(s) and the direction in which that gene is transcribed. Deletion of the right end of pBQ65 (Fig. 8) extending from the *Sst*I site (pBQ71) results in complete elimination of the 18-kDa polypeptide(s). Deletion of the centrally located sequences extending from the *Eco*RV site at 0.57 kilobase pairs to the *Pvu*II site at 1.7 kilobase pairs (pBQ69, Fig. 8) results in the loss of all three polypeptides and the appearance of a new polypeptide species migrating at approximately 15 kDa. This 15-kDa polypeptide probably is a truncated form of the 18-kDa

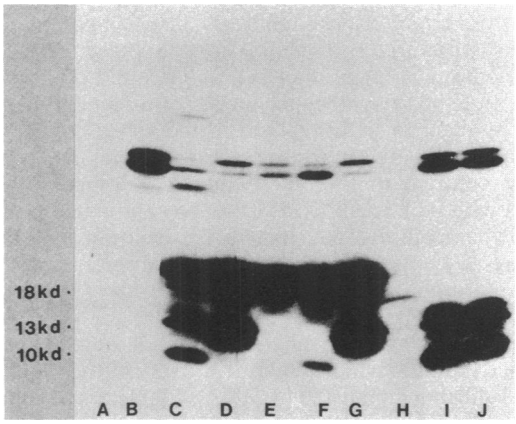


FIG. 9. Selective expression of polypeptides encoded by pBQ65 and its deletion derivatives in minicells. Purified minicells were labeled with [$^{35}$S]methionine for 60 min and then boiled for 5 min in final sample buffer. Polypeptides were resolved by electrophoresis on a 7 to 20% SDS–polyacrylamide gradient gel. Shown above is a fluorograph of the gel. Lanes: A, *E. coli* P678-54; B, P678-54 (pBR327); C, P678-54(pBQ65); D, P678-54(pBQ59); E, P678-54(pBQ67); F, P678-54(pBQ72); G, P678-54(68); H, P678-54 (pBQ69); I, P678-54(pBQ70); J, P678-54(pBQ71).

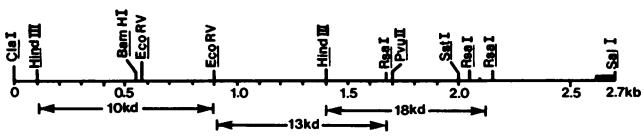


FIG. 10. Approximate locations of the polypeptide-coding sequences required for R-body synthesis. The arrows below the map of the cloned sequence in pBQ65 indicate the maximum limits of the genes coding for the three polypeptides known to be required for R-body synthesis.

polypeptide. Therefore, the regulatory sequences for the 18-kDa polypeptide(s) are located within several hundred base pairs to the right of the *Sst*I site (Fig. 8), and transcription proceeds leftward from these sequences. It should be noted that elimination of the regulatory sequence for the gene encoding the 18-kDa polypeptide(s) does not affect the expression of the other two gene sequences, strongly indicating that each gene is expressed as an independent transcriptional unit.

Analysis of polypeptides determined by the derivatives of pBQ65 and characterization of the organization of their genetic determinants strongly indicate that only two genes are involved in coding for the 13- and 18-kDa polypeptides. Examination of SDS-polyacrylamide and two-dimensional gels demonstrates that there is probably a group of two or three 18-kDa polypeptides and that the 13-kDa species can be separated into at least 12 different polypeptides (Fig. 7 and 9). However, when a deletion or rearrangement affects expression of one member of either species, all members of that polypeptide species are affected (Fig. 9). These results strongly support our supposition that posttranslational modification or processing of the 18- and 13-kDa polypeptide species accounts for the variations observed within each of these groups.

Fluorographs obtained by increased exposures of the gel used to generate Fig. 9 reveal limited polymerization in minicells containing deletion derivatives that are only unable to produce the 18-kDa polypeptide group (pBQ70 and pBQ71). Although a ladder profile is generated, these polypeptides migrate faster in SDS-polyacrylamide than the major ladder polypeptides associated with R-body synthesis. However, they do correlate with the minor ladder polypeptides produced in minicells containing the fully functional R body-coding plasmid pBQ65. These minor ladder polypeptides probably represent polymerization intermediates that are most likely precursors of the major ladder polypeptides. This observation leads us to suggest that the 10-kDa polypeptide may act before the 18-kDa polypeptide(s) in R-body assembly. However, the results described here do not provide any evidence to allow us to draw conclusions concerning the roles played by these polypeptides in posttranslational processing of the 18- and 13-kDa polypeptides.

## DISCUSSION

We have identified three genes, which appear to be independent transcriptional units, that are essential for the assembly and structure of type 51 R bodies. Considering that these genes appear to act as independent transcriptional units and that these genes are not located near any natural cloning vehicle promoters, it is likely that transcription of these genes is initiated from their natural promoters. These genes code for the synthesis of three species of polypeptides which migrate in SDS-polyacrylamide gels as if they possessed molecular masses of 10, 13, and 18 kDa.

Although the 18- and 13-kDa polypeptide species are each determined by a single gene, each species is composed of at least 3 and 12, respectively, distinct polypeptides. We can only conclude that the primary translation products of these genes are subject to posttranslational modification which involves alteration of apparent molecular mass and, in the case of the 13-kDa species, pI. The changes in apparent molecular mass are most likely due to removal of amino acids from the termini of the primary translation products. The change in pI observed in the 13-kDa species could be due to any number of modification reactions. Preliminary sequence analyses predict that the primary translation product of the gene coding for the 13-kDa species is a slightly basic polypeptide (D. Heruth and R. L. Quackenbush, unpublished observations). The modifications of the 13-kDa species may involve progressive acquisition of negative charges. However, the unusual feature of the modifications of the 13-kDa polypeptides, as reflected in changes in pI, is that at least five steps appear to be involved and that for each apparent step there is a stable (with respect to the other 13-kDa species) product that participates in polymerization. Thus, it is not clear whether each polypeptide represents the result of a single modification process on a 13-kDa primary translation product or a sequence of modification steps where the polypeptides possessing intermediate pIs represent intermediates in the process. Posttranslational processing of the 13-kDa polypeptide, which is undoubtedly the major structural protein of type 51 R bodies, possibly is necessary for cross-linking reactions that must occur in the polymerization and assembly processes involved in R-body maturation. In addition to posttranslational modification, there is a possibility that the slight variation in apparent molecular mass between the members of the 13-kDa pairs (Fig. 7) is due to the presence of alternate translational start sites. At present there is no evidence for this, except that a very weak possible ribosome binding site coupled with a possible ATG start codon is 24 base pairs upstream from the putative start codon for the 13-kDa polypeptide (D. Heruth and R. L. Quackenbush, unpublished observations).

The functions of the 10- and 18-kDa polypeptides are unclear. Pulse-chase experiments indicate that roles as structural components of R bodies are unlikely and that they probably act as catalysts in R-body assembly.

Since none of the methods known to dissociate noncovalently bonded protein complexes had any effect on the ladder profile, incorporation of R-body subunits into assembly intermediates probably involves the formation of covalent bonds. Ladderlike profiles of polypeptides similar to those obtained by polyacrylamide gel electrophoresis of purified R bodies, or minicells containing pBQ65, have been reported in association with two types of common pili produced by *E. coli* (10, 12). In both cases the polypeptide ladders were demonstrated to be the result of polymerization of pilus structural subunits. Although these observations support the conclusion that the polypeptide ladder associated with R-body synthesis consists of assembly intermediates, the mechanisms involved in the assembly of pili are probably not related to those responsible for R-body assembly. In contrast to R bodies, simple treatment of K88ab fimbria with SDS is sufficient to cause complete dissociation (10). Type 1 pili are resistant to dissociation by SDS, but they are susceptible to dissociation by exposure to low pH (8) or saturated guanidine hydrochloride (5). Reducing agents also exhibit no effect on the R body-associated ladder profile, indicating that the formation of disulfide bridges between subunits is not a factor in its stability. This observation is further supported by DNA sequence data indicating the absence of cysteine codons in the gene coding for the 13-kDa polypeptide (D. Heruth and R. L. Quackenbush, unpublished observations).

We have suggested that the 13-kDa polypeptides are the major structural subunits of R bodies. However, it appears that there is probably a large sequence (as much as 500 base pairs) between the coding sequences for the 13- and 10-kDa polypeptides that could code for a fourth polypeptide. If such a polypeptide exists, our failure to detect it may be due

to the absence of methionine residues. A second, smaller structural subunit would aid in explaining why the intervals in the ladder profile are irregular. However, it would not explain the anomalous migration of the ladder polypeptides with respect to the molecular weight standards in SDS-polyacrylamide gels. These polypeptides migrate slower in relation to molecular weight standards as the concentration of polyacrylamide increases. This observation cannot easily be explained, but it may indicate that the charge/mass ratio of the ladder polypeptides in the presence of SDS is greater than that normally observed with respect to SDS-polypeptide complexes. Since these observations are the opposite of what one would expect of glycoproteins (20), the possibility that the ladder polypeptides are glycoproteins is unlikely.

The results presented here cannot be used to suggest the nature of polypeptides involved in the assembly and structure of all types of R bodies. R bodies have been described in all species of *Caedibacter* (15) and in two species of *Pseudomonas* (7, 23). The R bodies produced by *Caedibacter* are of two types. Type 51, which occurs in *C. taeniospiralis*, unrolls (reversibly) from the inside when the pH is less than 6.5 and has acute angles at both termini. Type 7, which unrolls (irreversibly) from the outside, does not unroll in response to changes in pH and is blunt at the outer terminus of the rolled-up R body and acute at the inner terminus. The R bodies produced by the pseudomonads are also of two types, both of which differ from those produced by *Caedibacter* species. R bodies produced by *Pseudomonas avenae* are larger than those of *Caedibacter* species and are quite unusual in that the edges of the unrolled R body ribbon, running lengthwise, are not parallel (24). The termini of these R bodies form acute angles which extend all the way to the center of the R-body length. The R bodies of *Pseudomonas taeniospiralis* are smaller than those of *Caedibacter* species and exhibit characteristics of type 7 and type 51 R bodies (7). They resemble type 51 R bodies with respect to the fact that they unroll from the inside. However, they resemble type 7 R bodies in that they unroll irreversibly independent of pH changes and their outer termini are blunt. Hybridization studies involving the R body-coding sequence of pBQ65 with DNA preparations from *P. taeniospiralis* and various strains of type 7 R body-producing *Caedibacter* species failed to detect significant homologies (J. A. Kanabrocki, B. J. Cox, J. Lalucat, J. A. Dilts, and R. L. Quackenbush, submitted for publication). Type 7 R bodies and R bodies from both *Pseudomonas* species are not antigenically cross-reactive with antisera to type 51 R bodies derived from pBQ65 (Kanabrocki et al., submitted for publication; B. Wells and I. Gibson, personal communication).

On the other hand, the results presented here probably extend to all type 51 R bodies. Type 51 R bodies naturally occur only in *C. taeniospiralis*. The R body-coding plasmids in all strains of *C. taeniospiralis* exhibit no detectable differences with respect to distribution of restriction endonuclease recognition sites within their R body-coding sequences (16, 18). In addition to the R body-coding sequence of pKAP47, which is contained in pBQ65, the R body-coding sequence of pKAP116 has also been cloned and extensively mapped by restriction enzyme analysis. The restriction maps of these two R body-coding sequences are identical (F. Pond, S. Hodel, and R. L. Quackenbush, Abstr. Annu. Meet. Am. Soc. Microbiol. 1986, H134, p. 149). R bodies purified from five different strains of *C. taeniospiralis* were found to be immunologically cross-reactive with antisera to R bodies determined by pBQ65 (Kanabrocki et al., submitted). In fact, Western blots of the SDS-polyacrylamide gels containing purified R-body samples from these five strains of *C. taeniospiralis* revealed polypeptide profiles identical to that generated by the R-body sample obtained from *E. coli* P678-54(pBQ65) (Kanabrocki et al., submitted). We therefore conclude that the results presented here apply not only to assembly and structure of type 51 R bodies in *E. coli* but also to all strains of *C. taeniospiralis*.

## ACKNOWLEDGMENTS

We thank Gary Small and Judith Dilts for helpful discussions.

This research was supported by National Science Foundation grant PCM8113203 and Public Health Service Grant GM36293 from the National Institutes of Health.

## LITERATURE CITED

1. **Anderson, T. F., J. R. Preer, Jr., L. B. Preer, and M. Bray.** 1964. Studies on killing particles from *Paramecium*: the structure of refractile bodies from kappa particles. J. Microsc. **3**:395–402.
2. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7**:1513–1523.
3. **Clewell, D. B., and D. R. Helinski.** 1969. Supercoiled circular DNA-protein complex in *Escherichia coli*: purification and induced conversion to an open circular DNA form. Proc. Natl. Acad. Sci. USA **62**:1159–1166.
4. **Dilts, J. A., and R. L. Quackenbush.** 1986. A mutation in the R body-coding sequence destroys expression of the killer trait in *P. tetraurelia*. Science **232**:641–643.
5. **Eshdat, Y., F. J. Silverblatt, and N. Sharon.** 1981. Dissociation and reassembly of *Escherichia coli* type 1 pili. J. Bacteriol. **148**:308–314.
6. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
7. **Lalucat, J., O. Meyer, F. Meyer, R. Pares, and H. G. Schlegel.** 1979. R bodies in newly isolated free-living hydrogen-oxidizing bacteria. Arch. Microbiol. **121**:9–15.
8. **McMichael, J. C., and J. T. Ou.** 1979. Structure of common pili from *Escherichia coli*. J. Bacteriol. **138**:969–975.
9. **Meagher, R. R., R. C. Tait, M. Betlach, and H. W. Boyer.** 1977. Protein expression in *E. coli* minicells by recombinant plasmids. Cell **10**:521–536.
10. **Mooi, F. R., A. Wijfjes, and F. K. de Graaf.** 1983. Identification and characterization of precursors in the biosynthesis of the K88ab fimbria of *Escherichia coli*. J. Bacteriol. **154**:41–49.
11. **O'Farrell, P. H.** 1975. High resolution two-dimensional electrophoresis of proteins. J. Biol. Chem. **250**:4007–4021.
12. **Orndorff, P. O., and S. Falkow.** 1984. Organization and expression of genes responsible for type 1 piliation in *Escherichia coli*. J. Bacteriol. **159**:736–744.
13. **Pratt, C.** 1980. Kinetics and regulation of cell-free alkaline phosphatase synthesis. J. Bacteriol. **143**:1265–1274.
14. **Preer, J. R., Jr., and L. B. Preer.** 1967. Virus-like bodies in killer paramecia. Proc. Natl. Acad. Sci. USA **58**:1774–1781.
15. **Preer, J. R., Jr., L. B. Preer, and A. Jurand.** 1974. Kappa and other endosymbionts in *Paramecium aurelia*. Bacteriol. Rev. **38**:113–163.
16. **Quackenbush, R. L.** 1983. Plasmids from bacterial endosymbionts of hump-killer paramecia. Plasmid **9**:298–306.
17. **Quackenbush, R. L., and J. A. Burbach.** 1983. Cloning and expression of DNA sequences associated with the killer trait of *Paramecium tetraurelia* stock 47. Proc. Natl. Acad. Sci. USA **80**:250–254.
18. **Quackenbush, R. L., J. A. Dilts, and B. J. Cox.** 1986. Transposonlike elements in *Caedibacter taeniospiralis*. J. Bacteriol. **166**:349–352.

19. **Roozen, K. J., R. G. Fenwick, Jr., and R. Curtiss III.** 1971. Synthesis of ribonucleic acid and protein in plasmid-containing minicells of *Escherichia coli* K12. J. Bacteriol. **107**:21–33.
20. **Segrest, J. P., and R. L. Jackson.** 1972. Molecular weight determination of glycoproteins by polyacrylamide gel electrophoresis in sodium dodecyl sulfate. Methods Enzymol. **79**:237.
21. **Soberon, X., L. Covarrubias, and F. Bolivar.** 1980. Construction and characterization of new cloning vehicles. IV. Deletion derivatives of pBR322 and pBR325. Gene **9**:287–305.
22. **Talmadge, K., and W. Gilbert.** 1980. Construction of plasmid vectors with unique *Pst*I cloning sites in a single sequence coding region. Gene **12**:235–241.
23. **Wells, B., and R. W. Horne.** 1983. The ultrastructure of *Pseudomonas avenae* II. intracellular refractile (R body) structure. Micron Microsc. Acta **14**:329–344.
24. **Widmayer, D.** 1965. A non-killer resistant kappa and its bearing on the interpretation of kappa in *Paramecium aurelia*. Genetics **51**:613–623.